

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2003, 13:15:55 ; Search time 74 Seconds

(without alignments)
1623.730 Million cell updates/sec

Title: US-09-494-297-2

Perfect score: 3945

Sequence: 1 MKKTRFPKMLTNTQRYLS.....TAGISLGIWGIHTIRIKRD 757

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3945	100.0	762	23 ABP29047	Streptococcus poly
2	236.5	6.0	23	ABP29050	Streptococcus poly
3	181	4.6	905	20 AAM89421	Moraxella catarrha
4	180	4.6	597	21 AAY90257	Streptococcus equi
5	178	4.5	898	20 AAM89413	Moraxella catarrha
6	161	4.1	1416	20 AAY00211	Enterococcus faeca
7	161	4.1	23	ABP43430	E faecalis EF104 a
8	161	4.1	1416	24 ABU13709	Enterococcus faeca
9	161	4.1	1448	20 AAY00210	Enterococcus faeca

10	161	4.1	1448	23 ABP43429	E faecalis EF104 p
11	161	4.1	1448	24 ABU13708	Enterococcus faeca
12	159	4.0	1612	23 ABP47333	Listeria monocytog
13	156	4.0	1530	23 ABP47329	Listeria monocytog
14	155	3.9	832	23 ABP47336	Listeria monocytog
15	154	3.9	1849	18 AAM17900	Photobacterium lum
16	152.5	3.9	2013	23 ABP47751	Listeria monocytog
17	152	3.9	1849	19 AAM56573	Toxin TcdA11, enco
18	152	3.9	2516	18 AAM17899	Photobacterium lum
19	152	3.9	2516	19 AAM56572	Toxin TcdA, encode
20	152	3.9	2516	22 AAB72609	Photobacterium tcdA
21	152	3.9	2516	23 ABP32651	P. luminescens (W-
22	152	3.9	2517	22 AAB72611	Modified Phototrab
23	152	3.9	2537	22 AAB72614	TcdA toxin-zeln ER
24	150	3.8	1161	23 AAE22273	Streptococcal fibr
25	149	3.8	1315	20 AAY08642	S. aureus SdrD pro
26	149	3.8	1315	24 ABU18969	Pathogen specific
27	149	3.8	2032	20 AAY00238	Enterococcus faeca
28	149	3.8	2032	20 AAY00240	Enterococcus faeca
29	149	3.8	2032	20 AAY00242	Enterococcus faeca
30	149	3.8	2032	23 ABP43457	E faecalis EF123 p
31	149	3.8	2032	23 ABP43459	E faecalis EF124 p
32	149	3.8	2032	23 ABP43461	E faecalis EF125 p
33	149	3.8	2032	24 ABU13736	Enterococcus faeca
34	149	3.8	2032	24 ABU13738	Enterococcus faeca
35	149	3.8	2032	24 ABU13740	Enterococcus faeca
36	147	3.7	1349	22 AAY034402	Staphylococcus aur
37	147	3.7	1349	22 AAY034404	Staphylococcus aur
38	147	3.7	1563	22 ABB58432	Drosophila melanog
39	146	3.7	1185	13 AAR22675	Collagen binding p
40	145	3.7	1112	20 AAY08603	S. pyogenes SFRP-
41	144.5	3.7	940	23 ABB47334	Listeria monocytog
42	144.5	3.7	1092	19 AAM41602	Staphylococcus epi
43	144.5	3.7	2261	24 ABU18914	Pathogen specific
44	144.5	3.7	2283	24 ABP56876	Staphylococcus epi
45	144.5	3.7	2244	22 AAY037120	Staphylococcus aur

ALIGNMENTS

RESULT 1	ABP29047	standard. Protein; 762 AA.
ID	ABP29047	
XX	ABP29047;	
AC	02-JUL-2002 (first entry)	
DT		
XX		
DE	Streptococcus polypeptide SEQ ID NO 7270.	
XX		
KW	Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;	
KW	group A streptococcus; Streptococcus pyogenes; antibacterial;	
KW	antiinflammatory; infection; vaccine; meningitis; gene therapy.	
OS	Streptococcus pyogenes.	
XX		
PN	WO200234771-A2.	
XX		
PD	02-MAY-2002.	
XX		
PF	29-OCT-2001; 2001WO-GB04789.	
XX		
PR	27-OCT-2000; 2000GB-0026333.	
XX		
PR	24-NOV-2000; 2000GB-0028727.	
XX		
PR	07-MAR-2001; 2001GB-0005640.	
XX		
PA	(CHIR-) CHIRON SPA.	
XX	(GENO-) INST GENOMIC RES.	
XX	Telford J, Masiagnani V, Margarit Ros YI, Grandi G, Fraser C;	
PI	Tetelin H;	
XX		

DR WPI: 2002-352536/38.
 DR N-PSDB: ABN69678.
 XX New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 PS Claim 1; Page 3879; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 CC
 XX
 SQ Sequence 762 AA;

Query Match 100.0%; Score 3945; DB 23; Length 762;
 Best Local Similarity 100.0%; Pred. No. 2,4e-285;
 Matches 757; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKTREPKNLTINTQVLSKNSKRTVTLVGVFLMIFALVTSKVGAKTEGVLESSTPN 60
 DB 6 MKKTREPKNLTINTQVLSKNSKRTVTLVGVFLMIFALVTSKVGAKTEGVLESSTPN 65
 QY 61 AINPDSSEYRWYGYESYVGHPPYKQFVAHDLRVNLEGSRSYQVYCFNLKKAFLPGSD 120
 DB 66 AINPDSSEYRWYGYESYVGHPPYKQFVAHDLRVNLEGSRSYQVYCFNLKKAFLPGSD 125
 QY 121 SSYKRWYKKHDKISTKREDYAMSPRITGDELNOKLRVVMNGHQNANGMEGLEPINA 180
 DB 126 SSYKRWYKKHDKISTKREDYAMSPRITGDELNOKLRVVMNGHQNANGMEGLEPINA 185
 QY 181 RVTOEAWYVSDNAPISNPDESFRSESNLVSTQSLSKROALKOILIDNLTAKMKQY 240
 DB 186 RVTOEAWYVSDNAPISNPDESFRSESNLVSTQSLSKROALKOILIDNLTAKMKQY 245
 QY 241 PDDFQSLFSESEDKGDKYNGYNLSGLVPTKPTPGDPMPMPNOPTTSVLRKYAI 300
 DB 246 PDDFQSLFSESEDKGDKYNGYNLSGLVPTKPTPGDPMPMPNOPTTSVLRKYAI 305
 QY 301 GDYSKILEGATLQLTGNVNSFOARVSSNDIGERITLSGTYTLTFLNSPAGYSTAEP 360
 DB 306 GDYSKILEGATLQLTGNVNSFOARVSSNDIGERITLSGTYTLTFLNSPAGYSTAEP 365
 QY 361 TFFVEAGKYVTIIDKQIENPNKEIYEPYVEAVNDEEFSVLTTOYAKFYAKNNGS 420
 DB 366 TFFVEAGKYVTIIDKQIENPNKEIYEPYVEAVNDEEFSVLTTOYAKFYAKNNGS 425
 QY 421 SQVYVCFNADLKSPPSEDEGKTMTPTDFTTGEVKKYTHIAGRDLEKTYVKKPRDTPDFLK 480
 DB 426 SQVYVCFNADLKSPPSEDEGKTMTPTDFTTGEVKKYTHIAGRDLEKTYVKKPRDTPDFLK 485
 QY 481 HIKVIEKGYREGQALIEYSGLTETQLRATQALAIYFTDSAEIDKCKLKDYGFGDMND 540
 DB 486 HIKVIEKGYREGQALIEYSGLTETQLRATQALAIYFTDSAEIDKCKLKDYGFGDMND 545
 QY 541 SLTAIVAKIIVEYADNSNPOLTDLPFIPNNNKYQSLIGNOWHPEDLVDIRMEDKKEVI 600
 DB 546 SLTAIVAKIIVEYADNSNPOLTDLPFIPNNNKYQSLIGNOWHPEDLVDIRMEDKKEVI 605

QY 601 PVTHNLTRKTYVTGLADRRKDFEHEITELKNNKOLLISQVTKDKTNLEFKDGKATINLK 660
 DB 606 PVTHNLTRKTYVTGLADRRKDFEHEITELKNNKOLLISQVTKDKTNLEFKDGKATINLK 665
 QY 661 HGESLTQSLPEGSYLVKFTEDSRGYKXVNSQVNAATYSKGTSTDEFLAPENKEPV 720
 DB 666 HGESLTQSLPEGSYLVKFTEDSRGYKXVNSQVNAATYSKGTSTDEFLAPENKEPV 725
 QY 721 VPTGVDPKINGYALAIYIAGISLGIWGIHTRIRKHD 757
 DB 726 VPTGVDPKINGYALAIYIAGISLGIWGIHTRIRKHD 762
 RESULT 2
 ABP29050
 ID ABP29050 standard; Protein; 340 AA.
 XX
 AC ABP29050;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 7276.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus pyogenes.
 XX
 PN W0200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB04789.
 XX
 PR 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 XX
 DR WPI: 2002-352536/38.
 DR N-PSDB: ABN69681.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 PS Claim 1; Page 3880; 4525pp; English.

XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 CC
 XX
 SQ Sequence 340 AA;


```

Db      394 GNRRTGTAKSLLIDONTMTAPFVKELFSKKNPNP--DPN-----SDT 434
QY      307 LEGATLQLTGDNVNSFOARVSSNDI-----GERIELSDGTTLTELNSPAGYSTAEP 359
Db      435 LEGGFYSESDEL-----AGKFLSDNATFVFGSKRDKTTPVATKTVYFS--TFE--KP 487
QY      360 ITFVEAGKYVTIIDG---QIENPKETIPEYSVEAYNDF-----EEFSVLT 405
Db      488 STFVGNENEGISLIDGKLNDEVNNOJEDETVPVSNKEYEYKNGRPNKQFTKKNASVQ 547
QY      406 QNVA-----KFTY-----AK--KNSSQVYVCFNADLKSPPSDEGCKTMTPDF 449
Db      548 KNPAYFGOHDKFYFNGNYVDLSAKEANKLGVSDPTSNKSLILAKYPPA---KYSTNKV 603
QY      450 T-----GEVKYTHIAGRDLFKYTVKPRDTPDFELK---- 480
Db      604 TKIYLOAKDKPTAIAKSIDHISFGEVLINDNKGKPTNSFYVGGADVSTQDPSAGK 663
QY      481 -----HIKVIKREKREKGAIEYSGLTETQLRAATQLAIFYFT----- 519
Db      664 FTYNGELWAGYLTQKKDKGYSKDEDTIKQKGLK-----YILTKEPIQDDDDDD 712
QY      520 DSALDLDKDKD--YHGEGDAMDSTLAVAKILVEYAQDSNPQLTDLDFIIPNNKKYOSLI 578
Db      713 DSLTASDSDODNTHGDDDL-----IASDSDODDITDGDSDDL--GDADDDAA 761
QY      579 GTOWHPEDLDVLIIMEDKKEYIPIY---THNLT-----RKVTGLADRT 620
Db      762 GKUYHAQN-----IRPEENKTLPIINEPTHEKTFALDGNKAKAFVNEEDTNSLTGLK 817
QY      621 KDEHELELNKNNKQELLISQVTKDRTNLEFKDGKATINLKHGESLTQGLPEGSYLVKE 680
Db      818 GDIFYDI--KNGKIDGTGTAKADVPNVRREVQ---NNOGG-----GLYINIKD 861
QY      681 TDSGCKYKYNQSOEVA 696
Db      862 IDVKGQFGTNGEELA 877

RESULT 4
ID AAY90257 standard; Protein: 597 AA.
AC AAY90257;
DT 19-SEP-2000 (first entry)
DE Streptococcus equi fibronectin binding protein, FN2.
KW Fibronectin binding protein; SFS; vaccine; horse; strangles; therapy;
KM equine upper respiratory tract disease; S. equi infection; FN2.
XX OS Streptococcus equi.
XX PN MO200037496-A1.
XX PD 29-JUN-2000.
XX PF 21-DEC-1999; 99WO-SE02448.
XX PR 22-DEC-1998; 98SE-0004491.
XX PA (GUSS/) GUSS B.
XX PA (LIND/) LINDMARK H.
XX PA (JACO/) JACOBSSON K.
XX PA (FRYK/) FRYKBERG L.
XX PI Guss B, Lindmark H, Jacobsson K, Frykberg L;
DR WPI; 2000-442641/38.
DR N-PSDB; AAA30874.
XX

```

```

PT      New protein useful for preparation of vaccines for treatment of
PT      strangles caused by Streptococcus equi infection, is able to bind to
PT      mammalian fibronectin -
XX      Disclosure; Page 17a-17b; 34pp; English.
XX      This sequence represents the Streptococcus equi fibronectin binding
CC      protein, FN2. The FN2 sequence was used to isolate the S. equi
CC      fibronectin binding protein of the invention, designated SFS. SFS binds
CC      specifically to mammalian fibronectin or its analogues or fragments. The
CC      protein, its analogues or fragments may be used for the preparation of a
CC      vaccine that protects horses against strangles (a world-wide distributed
CC      and serious disease of the equine upper respiratory tract) caused by
CC      S. equi infection. The antibody and/or antiserum may also be used for the
CC      prophylactic or therapeutic treatment of S. equi infection in mammal,
CC      especially horses. The use of vaccines containing the fibronectin binding
CC      protein provides a more effective protection against S. equi infections,
CC      with fewer side effects.
XX      Sequence 597 AA;
SQ
Query Match 4.6%; Score 180; DB 21; Length 597;
Best Local Similarity 20.8%; Pred. No. 0.00016;
Matches 149; Conservative 102; Mismatches 265; Indels 202; Gaps 34;
QY      68 SEYRWYGESYVR-GHPYKQFVYADLRVNLGSSRSYQVCFNLKKAPFLGSPSSYKRW 136
Db      32 AEQLYYQWNGTQRSSPYF--LVYSPKNAPKRELKDEYVYVCFNKLIYWDQMSIYSNF 89
QY      127 -----YKNDGISTKFEDYAMSRLTIGDELNOKLRVAVYNGHQMANGINEG- 173
Db      90 NDIRSPNDLRYVEKKGIDYGYDIFKQYAPDKKIDISALVAVLSNGYPTNKSQLSY 149
QY      174 -LEPLNLRVTOEAVWYYSNAPISNPDESFRSESNLSTQSLSMRQALQOLIDPNL 232
Db      150 HLNNDSSRKVTQLAIMTFSD---SLTKEFLKDTGGTNL-----NDMEKALDFILSKGE 200
QY      233 ATKMPKQVPDPFOLSIPESEDKDKYKQYONLLSGGLVPTKPPGDPMPMPNOQTTS 292
Db      201 DSKL-KSEQSNYSLDIYVQSGHDHMKDYNLLGSLTILPEKLPK----- 245
QY      293 VLIRKVAIGDYSKLBEATQLTGDNVNSFOARVSSNDIGERIELSDGTTLTELNSPA 352
Db      246 -----QLGGS-----GHNGSLGLEGGSSGSQETNE--DGKKGL--IGFHG 284
QY      353 GYSIAE---PIF-FKVEAGKYVTIIDGKQIENPKETIPEYSVEAYNDEEFSVLTTON 407
Db      285 GLSGSEGRDPLGLKGEAGAPD-----PQKPNDL---QGLEGN-----SPIVEON 330
QY      408 YAKFYAKNKNSSQVYCFNADLKSPPSDEGCKTMTPDFTGEVAKYTHIAC-----R 461
Db      331 YGS---TEGYHGOSGILE--ETEDTNPGIILGSGNVE--THEDTRNPHIMGIGGLAG 383
QY      462 DLEKTYPKR-----DTDPDTFLKHKKYIEKGYRREGOALIESGTLTQLRAATQ 512
Db      384 ESGETTPKPGQTGGQGVETTEDT-----OKGMSGGSGGTTIESENKTKPE 429
QY      513 LAIYFTDSALDKDKLDYHGDMNDSTLAVAKILVEYAQDSNPQLTDLDFIIPNNN 572
Db      430 VMIGGQQTLETETEDTK--GMSGGSGGITE-----SEDTKREV----- 467
QY      573 KYOSLIGTOWHPEDLDVLIIMEDKKEYIPVTHNLTLKTYVTGLAGRTKPFHEIELKN 632
Db      468 -----MIGGQ-----GQIIDESEN--TQSGMSGSGGDTT-----VIEDTK 500
QY      633 KOELLISQVTKDNTNLEFKDGKATINKKHESLTQGL--PEGSYLVKREDSGKYKYV 690
Db      501 KSEIT--IGGQQLIDSESD-----TQPMGSGGSGTIVETDKTPTRPKP 545
QY      691 NSQEVANATVSKTIGTSDEFIAPENKPEVPTGVQDQ---INGLALIVYAGISL 743
Db      546 APAPIYV-----DE-----KPNKGTHLPQTSMDKQLTSLISGMSLVLVCLSL 589
XX

```

```
RESULT 5
AA089413
ID AA089413 standard; Protein: 898 AA.
XX
AC AA089413:
XX
DT 21-JUN-1999 (first entry)
XX
DE Moraxella catarrhalis lactoferrin binding protein 2 (Lbp2).
XX
KW Lactoferrin receptor; lactoferrin binding protein; Lbp2;
KW Lbp2 gene; infection; otitis media; sinusitis; conjunctivitis;
KW pneumonia; bronchitis; tracheitis; emphysema; diagnosis; therapy;
KW vaccine; Branhamella catarrhalis.
XX
OS Moraxella catarrhalis.
XX
FH Key Location/Qualifiers
FT Misc-difference 632
FT Region /note="encoded by AAR"
FT /note="conserved epitope"
FT 430..435
XX
PN M09855606-A2.
XX
PD 10-DEC-1998.
XX
PF 02-JUN-1998; 98MO-CA00544.
XX
PR 08-MAY-1998; 98US-0074658.
PR 03-JUN-1997; 97US-0867941.
XX
PA (CONN-) CONNACHT LAB LTD.
XX
PI Du R, Klein MH, Loosmore SM, Wang Q, Yang Y;
XX
DR WPI: 1999-070266/06.
XX
DR N-PSDB: AAV82019.
XX
PT Lactoferrin receptor genes from Moraxella, especially M. catarrhalis
PT useful to diagnose Moraxella infection e.g. to detect otitis media
PT due to M. catarrhalis infection and to immunise against such
PT infections
XX
PS Claim 8; Fig 2; 202pp; English.
XX
CC This protein comprises lactoferrin binding protein 2 (Lbp2) of
CC Moraxella catarrhalis (Branhamella catarrhalis) 4223. It is
CC encoded by the lbp2 gene of the lactoferrin receptor (lfr) locus
CC (see AAV72019) identified in the M. catarrhalis 4223 genome.
CC Immunogenic compositions, including vaccines, based upon expressed
CC recombinant lbp1 and/or lbp2 and/or ORF3 proteins (see AAV89413-21),
CC portions of these, or their analogues, can be prepared for
CC prevention of diseases caused by Moraxella. M. catarrhalis is a
CC causative agent of otitis media and has been associated with
CC sinusitis, conjunctivitis and inflammatory diseases of the lower
CC respiratory tract, such as pneumonia, chronic bronchitis,
CC tracheitis and emphysema.
XX
SQ Sequence 898 AA;
Query Match 4.5%; Score 178; DB 20; Length 898;
Best Local Similarity 18.6%; Pred. No. 0.00043;
Matches 171; Conservative 118; Mismatches 305; Indels 324; Gaps 42;
OY 4 TRPNNKTLNTFOVLNSKSRFTVTLVGVFLMTFALVTSVNGAKTVGLVSSNPNNIN 63
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 79 TTPDNGDNNOITQ-----AQKTAAGAAGFFVM-----GKTRDTPSKN-D 115
OY 64 PDSSE--YRWYG--YESYVGRHP-----YYKQFVAHDLRYNL 98
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 116 PDYSNDLVQOMQKGLYIGDAHRPDGIGTGKLNQPIITANDIKPLDYFVKFPALSDLHDS 175
```

```
OY 99 EGSRSYQVYCFNKKAPFLGSDSSVK--WYKKHGISTKPEDYAMSP-----RIT 147
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 176 ERHR-FDPKRLNTJIKVGYGNLTPPSKNNTYINHQADNKNRPVDYENIRFGYLELQ 234
OY 148 GDELNOK-----LRVWYNGHPONANGIMEGLEPLN----- 178
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 235 GSSLTKKADPNDKDRIPKPMPIIFHYG--EMASSQLPSAGKRFYNTGNMLYLDVKKRP 292
OY 179 AIRVTOEAWWYSDNAPISNPDESFRSESNLVSQSL- 219
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 293 ALSASDDRVGYVLNASKSN-----EDDVYSAHIYINGFQYKHTPATYQVDFDTN 343
OY 220 -MROALKOLIPNLTAKPKOYPPDF-----QLSTFESDK--GDKRYNGYONLSGG 269
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 344 SLTGKLSYDNPNOATGAKYIKSQFDTKRVNETDVQIDAKINGNPFVGTASLVNEN 403
OY 270 L-----VPTKPPTPGDPMPNPQPTQTSVLIRKVAIGDYSKLLEGATLQLGDVNS 321
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 404 TETAPFTELFSKANPNP--NRP-----SDLEGGFYGESDEL-- 442
OY 322 FOARVSSND-----IGRIELSDGYTLTELNSPAGYSIAEPTFKVEAGKYTTIDG 375
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 443 --AGFLSNDNASYVFGKRDKTDKPVATKVYFSAGFE--KPSTSFVDNETIGRIINS 498
OY 376 KOIENPKIYE-----PYSVAIYND-----EERSVLTION-----YAKFYAK 415
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 499 KRLNDVAVEKIDNEDIPTSERYDEFPWGEKKAETKRVSSSTAQAVPAYFGOHDKFYE- 556
OY 416 NKNSSQVYVCFNADLKSPDSEDEGKMTPTDFTTGEV----- 453
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 557 --NGNYDLSASSVDKILAPADAVKANOSIKREKYPATLNKNQYTAIYLOEAKOKKPYTA 614
OY 454 ---KYTHIA-GRDLF-----KYTVKPRDIDPTFLHKKVI----- 486
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 615 IRASYOHIISFGETLYNDANOTPRSYFVQGGRADTSTLPKAGFTYNGIMAGYLQKK 674
OY 487 EKGYREKQALIEYS-----LTETQLRAATQLATIFYT-----DSALDLDKLDYNGF 535
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 675 DKGSNNBETIKKGHODYLLTED-----FTPEDDDDLASDSQDDADAG 722
OY 536 GDMNDSTLAVKILIVEYQDSNPOLTDLDFIPNNKYOSLIGTQWHPEDLVDIRMED 595
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 723 DDL-----IASDSQDDADAGDDSDDL--GDGADDAAGKAVYHAGN-----IRPEF 767
OY 596 KKEVIIV--THNLT-----RKTVYGLAGORTKDFHEIELKNNKQELL 637
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 768 ENKYLIPINEPHEKTFALDGKNAKAFVDVDFDNTSLTKLNDERGDIYFDI--KNGKIDGT 825
OY 638 SQTAKTDKTNLEFKDGKATITLKHGESLTLQGLPEGYLVKPEDSEGYKKVANSQEA- 696
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 826 GFTKADVPNRYREEG--NNQG-----GFLYNIKIDIVKQGFQGTGEEELAG 871
OY 697 -----NATVSKTG 704
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 872 QLQYDKGDIINDTAKEAG 889
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 6
AA000211
ID AA000211 standard; Protein: 1416 AA.
XX
AC AA000211;
XX
DT 20-APR-1999 (first entry)
XX
DE Enterococcus faecalis antigenic polypeptide fragment EF104.
XX
KW Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
KW detection; attenuation; antigenic.
XX
OS Enterococcus faecalis.
XX
```

PN W09850554-A2.
 XX
 PD 12-NOV-1998.
 XX
 PF 04-MAY-1998; 98WO-US08959.
 XX
 PR 14-NOV-1997; 97US-0066009.
 PR 06-MAY-1997; 97US-0044031.
 PR 16-MAY-1997; 97US-0046655.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Bailey C, Choi GH, Hromockyj A, Kunsch CA;
 XX
 DR WPI; 1999-070095/06.
 DR N-PSDB; AAX20201.
 XX
 PT New isolated Enterococcus faecalis polynucleotides - used to develop
 PT products for the detection of Enterococcus and for use in vaccines
 PT for prevention or attenuation of Enterococcus infection
 PS
 PS Claim 9; Page 203; 301pp; English.
 XX
 CC The present sequence represents an antigenic polypeptide fragment
 CC isolated from Enterococcus faecalis. The present invention describes
 CC genes, proteins and antigenic polypeptides isolated from E. faecalis.
 CC The proteins can be used in vaccines for preventing or attenuating an
 CC infection caused by a member of the Enterococcus genus in an animal.
 CC They can also be used for detecting Enterococcus antibodies in a sample.
 CC The nucleotide sequences can be used for detecting Enterococcus nucleic
 CC acids. Products from the present invention can also be used for
 CC screening compounds to identify agonists and antagonists of E. faecalis
 CC protein activity.
 CC
 SO Sequence 1416 AA;
 Query Match 4.1%; Score 161; DB 20; Length 1416;
 Best Local Similarity 21.1%; Pred. No. 0.017;
 Matches 159; Conservative 104; Mismatches 326; Indels 166; Gaps 40;
 QY 39 ALVTSMVGAKTIVGLVES---STPNAINPDSSEYRMVGESEYVRGHPYKKORVAHDLR 95
 DB 484 SLSTPIVIGPKAIQVSDOTIEPIISVNP-LNLETAMGNTDO-----NGAYSSR 531
 QY 96 --VNEBGSRSYQYVYCNLKKAF-----LGSDDSVKKYKKHDSITKFEEDYAMSPRTGD 149
 DB 532 TTYSVWGSKKEPIQNLKIKHNYLSLRATKEIYYK---LGT---DYTTPTSDGS 584
 QY 150 ELNOKLRAVMYNGHPONANGIMEGL-----PLNAIRVTOEAVWYSDNAPISNPDESFK 204
 DB 585 VIFETTPITNEIQIPGFNVVPDPLPKDSIPVDITPITWSAGLTPVDTTVT---NSK 641
 QY 205 RESESNLVSTQSLSLRKQALKOLIDPNLTKMKQVDPDQSLFSESDGDKYNGCYON 264
 DB 642 RGSERTLOSCKNOFLVARNDSFSLSVRKIPAGA--DVLFDIYDVSN--DOVDSIYPO 697
 QY 265 LLSGLVPTKPTPTGDPMP--PNOPQTSVLI-----RKVAYIGYSKL---LEGATQ 313
 DB 698 YMRGQYFPKPMPPNSPGYPTTFEDNTNSYTFDFPKTKNRITII-EKKNANGIDPTLY 756
 QY 314 LIG-----DNVSFQARVSSNDIGERIELSDGYTLTLELNSPAGYSIAEPITFVEAG 367
 DB 757 ITGTAKPEQSNNEGASASVQON---EALDIISAT-----QAANPILKNVTKT 801
 QY 368 KYVT-IIDCK-QIENPKIEIVPEVEANDFEESVLTQNYAKFYAKNNGSSQVY 424
 DB 802 TVTTKNIDNKTNHNKMPTELETPKGTNMOIDNLSITV-----KGVPEDA 846
 QY 425 YCFNADLKSPDSEDDGKTPTPFTTGE---VKYTHIAGRLDKRYTVKPRDTPDFLKH 481
 DB 847 Y-----SLEKTTNGAKVIFKRYTLENTIIEYNVSANAGQIYETITDSETLQMSA 899
 QY 482 IKKVI-----EKGYREKGAILEYSGLTETQLAATOLAIYVFTD---SAL 524

DB 900 SKKKVTTAPITLKESGDEAGIYYLAFAFYTHNVEDENQALAKVSFEILDVNTHTATER 959
 QY 525 DKDKLADYHGFQM-RNSTLAIVAKILVEYAQDSNPOLT-----DLDFRIP----- 569
 DB 960 TTDEKQGYSFDAIMTGDYTLRTVNPQESVDEE--YLGGKAIKLYKGDMQKILPLTKTI 1017
 QY 570 NNNKYO-----SLIGTQWHEPEDIVLIDIMEDKK-EVYIPYHNLTKRYTGLAGDRTKDF 623
 DB 1018 DHSRLQYKDSITIVGDSMKPEE--NFVSATDKTGQDVPEF-----KITVSGYDNKKAGV 1070
 QY 624 HFEIELKNNKQELLQVTKTKTNLEFKDGKATINLKHGSELTGQLPEGYSLYKEDS 683
 DB 1071 YPIIYSDGKEETFAVYTVKPDQSKLEKVD--TTIYV--GDSWK-----PE--DNFVSATDK 1120
 QY 684 EGYK---KNSQEVANATYSKGITSDELAEEN 715
 DB 1121 TGQDVPEKIDVQGYVN--VDKITG---DYEIVKVN 1150
 RESULT 7
 ID ABP43430 standard; Protein; 1416 AA.
 XX
 AC ABP43430;
 XX
 DT 05-AUG-2002 (first entry)
 XX
 DE E faecalis EF104 antigenic fragment.
 XX
 KW Enterococcus; vaccine; gastrointestinal disease; diagnosis; antidiotic.
 XX
 OS Enterococcus faecalis.
 XX
 PN US2002045737-A1.
 XX
 PD 18-APR-2002.
 XX
 PF 04-MAY-1998; 98US-0071035.
 XX
 PR 04-MAY-1998; 98US-0071035.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;
 XX
 DR WPI; 2002-425450/45.
 DR N-PSDB; ABN98186.
 XX
 PT New genes and polypeptides from Enterococcus faecalis, useful as
 PT vaccines for preventing, treating or attenuating an infection caused by
 PT a member of the Enterococcus genus in an animal, particularly E.
 PT faecalis
 PS
 PS Claim 9; Page 185; 255pp; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of polypeptides from Enterococcus faecalis. The proteins can be
 CC used as vaccines for preventing or attenuating an infection caused by a
 CC member of the Enterococcus genus in an animal, particularly E. faecalis.
 CC The polynucleotide is also useful for preventing or treating E. faecalis
 CC infection. The present sequence is a protein of the invention.
 CC
 SO Sequence 1416 AA;
 Query Match 4.1%; Score 161; DB 23; Length 1416;
 Best Local Similarity 21.1%; Pred. No. 0.017;
 Matches 159; Conservative 104; Mismatches 326; Indels 166; Gaps 40;
 QY 39 ALVTSMVGAKTIVGLVES---STPNAINPDSSEYRMVGESEYVRGHPYKKORVAHDLR 95
 DB 484 SLSTPIVIGPKAIQVSDOTIEPIISVNP-LNLETAMGNTDO-----NGAYSSR 531

OY	96	-VLEGSRSYQVYCFNLKKAFF-----LGSOSVKKWKKKHHGDISTREKEDAMSRLTGD	149
Db	532	TTYSVMGSKREKPIQNLLEIKHHPNLSLRATKEIYFYK-----LGT--DYTVPTSDGS	584
OY	150	ELNOKLRVWVNGHPONANGIMEGLE-----PLAIRVTOGAWVYYSNAPISNDESFK	204
Db	585	VIKRTPTITNIOIPIGENTYPPDLPRDKSLPVDIPIITMAEGLTPDITVTY---NSK	641
OY	205	RESESNLVSTISQLSLMRQALQOLIDNPILATKMPKOVPPDFQLSIFESBDKDKYKGYON	264
Db	642	RGSERTLQSSKNQGLVNARNDSFDLSVTRKIPAGA--DYLEDIYDVSN--DQYDSIYQP	697
OY	265	LLSGGLVPTKPTPTGGDPMF--PMPQPTSVLT-----KKYALGDYSKL--LEGATLQ	313
Db	698	YMDRGQYFDDKFMPTNSPGYPTITPEDENSTYTFDFGKTKNKRYII-EYKNANGWIDVPTLY	756
OY	314	LTG-----DVNVSFOARVSSNDIGERIELESDGTPTTELNSPAGSIAPIPFKVGAG	367
Db	757	ITGTAKEPQSNNGSASVSQYQ---EALDILSAT-----QAANPTLKNVTKT	801
OY	368	KVYT-IIDGK--QIENPKELIEVEYSVLAANDPEEFSVLTQNTAKFYAAKKNKSSQVY	424
Db	802	TVTKKNIDNKTNRVKNPTIELTPRGTTNAQIDLSITV-----KGVPEDA	846
OY	425	YCFNADLKSPDSDSDGKGTMPDPTTGE---VKTHLAGRLKITYKPRDTPDPTLKH	481
Db	847	Y-----SLEKTTNGAKVIFKDDTLTENITIEYNTVSANGOJYETTTIDSELTLMQMSA	899
OY	482	IKKYV-----EKGYRRKGOAIEYSGLEPQLRAATQALAIYFID---SAEL	524
Db	900	SKKVVTTAPITLAKSEGDAGETIYLATFTFTTHVEDENQALAKVSFELIDNVTHTATEF	959
OY	525	DKDKLDKYHGEGDM-NDSTLAVAKILVEYAODSNPOLT-----DLDFEIP-----	569
Db	960	TTDEKQGYSPFAIMTGDYTLRVTVNPQEYSVDEE-YLTGAIKILYKGDNDQKPIPTKI	1017
OY	570	NNNKYQ-----SLIGTQWHPEDLVDIIRMEDK-EVLPVTHNLTKTGTVLGADRTKDF	623
Db	1018	DHSHLQYKDSPTIYVGDMSWKPEE--NFVSATDKTGQDVFE-----KIVSGQVDNXXKAGV	1070
OY	624	HFEIELKNNKQELLQYVKTDKTYLPERDGAATINLKHGESLTLQGLPEGYSYLKVEDS	683
Db	1071	YPIIYSDGEKETAIVYVVKPQDSKLEVKD--TIIYV--GDSWK---PE--DNFVSATDK	1120
OY	684	EGYKV---KYNSQEVANATVSKTGITSDETLAFEN	715
Db	1121	TGQDVPEPKIDVGSTVN--VDKIG---DYELIVKN	1150
RESULT 8			
ABU13709			
ID	ABU13709	standard; Protein: 1416	AA.
AC	ABU13709;		
DT	26-FEB-2003	(first entry)	
DE	Enterococcus faecalis EF040	polypeptide #202.	
XX	Enterococcus faecalis.		
XX	US6448043-B1.		
XX	10-SEP-2002.		
XX	04-MAY-1998;	98US-0071035.	
XX	06-MAY-1997;	97US-044031P.	
XX	16-MAY-1997;	97US-046655P.	
XX	14-NOV-1997;	97US-066009P.	
XX	14-NOV-1997;	97US-066009P.	

Query Match	Best Local Similarity	21.1%, Pred. No. 0.017;	Matches 159;	Conservative 104;	Mismatches 326;	Indels 166;	Gaps 40;
39	ALVYMSGAKVPEGLVES---	STFNAINPDSSSEYRWGYESYRGHPHYKQFRYAHDLR	95				
484	SLSTPVGIPGNKAIQVDSQDYIEPTISVNP-	LNALFTAMGNYQ-----	NCAYSSR	531			
96	--VWLSESRSYQVYCNFLNKAFP---	LGSDDSYAKWKKKHGDISTKEDYAMSRLTGD	149				
532	TTVSVMSSKEKPIQNLKVKHNPYLSIRAKKEIFYRK	-----LGT---	DYTVPTPSGS	584			
150	ELNQLKRAVMYNGHPQANANGIMEGLE---	PLNAIRYQDAVMYYSQDAPISNPDESFK	204				
585	VIRKPTPTITNEIQIPIGFNVYVDSLPKDKSIPIVDT	PIITMGAEGLTPTVT---	NSK	641			
205	RESSSNLVTSQSLMROALKQOLIDPNLATMPKQVPPD	FDPLSIFESDKDKTKNGKQYN	264				
642	RGSERTLOSSKNOGLVNAANDSPSLSVKRTIPIAGA-	DVLFEDIYVSN--	DOVDSIYPQ	697			
265	LLSGGLVPTKRPPTGGDDPMP-	PNQOPQTSVLI-----	RKYAIGDYSKL--	REGATLQ	313		
698	YMDGQYFDRKPMTNPSQCPYPTITTDENTNSTPT	DFGKTAKRII--	EYKANGMWDIVPLY	756			
314	LTG-----	DNVNSFOARVFSSNDIGERIELSDQYTLTLELNSPAGYSIAPITFEYAG	367				
757	ITGAKPEQOSNNNGSASVSQYN--	EALDILSAT-----	QAMPPLKNTYKT	801			
368	KVYI--IIDGK--	QIENPKKEIVEYISVEAYWDEEFVSYLTQNTAKFYAKKNGSSQVY	424				
802	TVTTRKNIDNKRHRKNPITIELTPKGTMAQIDLSITV-		RGVPEDA	846			
425	YCFNADLKSPDSESDGKTKMPDPTTGC---	VKTYHLAGRDLFKYTVKPRDTDDTFLKH	481				
847	Y-----	SLEKTINGAKVIRKDYTLTLENTITIEYNTYSANAGQIYTETTIDSEFLNOMSA	899				
482	IKKVI-----	EKGREKGAQIEYSGLTQLRAATQALAIYFTD---	SAEL	524			
900	SKKVVTAAPITLTKFSEGDAGIEIVYLAATATFT	HNVEDBNQAIKAVSFELDNVHTATF	959				
525	DKDKLAKDYHGFQDM--	NDSTLAVNAKILVEYADDSNPOLT-----	DLDFLR-----	569			
960	TTPEKGQYSPDAMTGTGYTLTAVTNVPEYSVDEE-	YLTGKAIKLKLVNGDNQKILPLTKI	1017				

```

QY 570 NNNKYYQ-----SLIGTQWHEBDLVDIIRMEDK-EVLPVTHNLTLRKTVTGLADGRTRDF 623
DB 1018 DHSRLQYKDSSTIYVGDSSMKPEE--NFVSATDKGQDVPE-----KITVSGQVONXXAGV 1070
QY 624 HFEIELKNNKQELLSSQVVKDKTNLEFKDKATINLKHGSLTLQGLPEGSILYVKETDS 683
DB 1071 YPIIYSDEGKEETAYVVKPDQSKLEVKD--TTIYV--GDSWK---PE--DNFVSATDK 1120
QY 684 EGYKV---KVNSEVANATVSKTGITSDETLAFEN 715
DB 1121 TGDVPEPEKIDVQGTVN--VDKIG---DYELVYKN 1150

RESULT 9
AA00210
ID AA00210 standard; Protein; 1448 AA.
AC AA00210;
XX
XX 20-APR-1999 (first entry)
XX
XX Enterococcus faecalis protein EF104.
DE
XX Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
KM detection; attenuation; antigenic.
XX
XX Enterococcus faecalis.
OS
XX W09850554-A2.
PN
XX 12-NOV-1998.
PD
XX
XX 04-MAY-1998; 98WO-US08959.
PE
XX 14-NOV-1997; 97US-0066009.
PR 06-MAY-1997; 97US-0044031.
PR 16-MAY-1997; 97US-0046655.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Bailey C, Choi GH, Hromockyj A, Kunsch CA;
PI
XX WPI: 1999-070095/06.
XX N-PSDB: AAX20200.
DR
XX
XX New isolated Enterococcus faecalis polynucleotides - used to develop
PT products for the detection of Enterococcus and for use in vaccines
PT for prevention or attenuation of Enterococcus infection
XX
XX Claim 9; Page 201; 301pp; English.
PS
XX
XX The present sequence represents a protein isolated from
CC Enterococcus faecalis. The present invention describes genes, proteins
CC and antigenic polypeptides isolated from E. faecalis. The proteins can
CC be used in vaccines for preventing or attenuating an infection caused
CC by a member of the Enterococcus genus in an animal. They can also be
CC used for detecting Enterococcus antibodies in a sample. The nucleotide
CC sequences can be used for detecting Enterococcus nucleic acids.
CC Compounds from the present invention can also be used for screening
CC compounds to identify agonists and antagonists of E. faecalis protein
CC activity.
CC
XX
XX Sequence 1448 AA;
SQ

Query Match 4.1%; Score 161; DB 20; Length 1448;
Best Local Similarity 21.1%; Pred. No. 0.017;
Matches 159; Conservative 104; Mismatches 326; Indels 166; Gaps 40;
QY 39 ALVTSMYGAKTVEGLVES---STFNAINPDSSSEFRYGYESYRGHPYKQFVARDLR 95
DB 511 SLSPVYGVKKAIOVSDQYIEPISVNP-LNAETLNGANTDQ-----NGATSSR 558
QY 96 -VNLEGSRSYQVYCFNLKRAF--LGSDDSVKWKYKKHKGISTKFEDYAMSPRITGD 149

```

```

DB 559 TTVSVMSKSEKPEIQLNLEIKYKHPNYVLSLRATKEIFYK-----LGT---DYTVPTSDGS 611
QY 150 ELNOKLRAVYVNGHPQANANGMEGL-----PLNAIVTQEAWVYSDNAPISMPDSFK 204
DB 612 VIKETPTITNEIQIPIGFNNVPPDSLPRKDISPVDTITPITMSAGELPVDVTVT--NSK 668
QY 205 RESESNLVSTQSLSLMRALKQLIDPMLATKMPKQVDPDDQLSIFSEEDKGDYKNKQYQN 264
DB 669 RGSERTIQLSSKNQGLVARNDSPDLSVRKIRPGA--DVLFDIYDSN--DQYDNLTPQ 724
QY 265 LLSGGLVPTKPPPGDPPM--PNQPTSVLI-----RKVAIGDYSKL---LEGATLO 313
DB 725 YWDRGQYFDKPMTPNSGPGYPTITFEDENTNSYTFDFGKTKRXYI--EYKNANGMIDVPTLY 783
QY 314 LTC-----DNVNSFOARFVSSNDIGRIELSDGTLYLTLSNPAGYSIAEPTFKYEAG 367
DB 784 ITGTAKEPQSNNGSASVSYQN--BALDLSAT-----QANPILKANVTKT 828
QY 368 KVPV-IIDGR--QIENPNKEIVEPYVEAYVNDPEEFSVLTONYAKFYAKNKGSSQV 424
DB 829 TVTTKNIDNTHRYKNFTIELTPKGTINQIDLSITV-----KGVEDA 873
QY 425 YCFENADLKSPDSEFGSKTMTPDFTGE--VKYTHIAGDLEKYVAKPDDPDTFLKH 481
DB 874 Y-----SLEKTINGAKVIFKDYTLTENITIEYNTVSAWAGQIYETETIDSETLNMMSA 926
QY 482 IKKYI-----EKGYREKGOAIEYSGLTETOLRAAOLAIYFTD-----SAEL 524
DB 927 SKRKVTTAPITLTKFSEBDARGIYVLAATATYTHNVEDENQAIKAVSELDINDTHATFEF 986
QY 525 DKDLKDYHGFQDM-NDSTLAVAKILVEYADSNPQLT-----DLDFETP----- 569
DB 987 TTDEKGGYSPDAIMTGDYTLRVTNVPOEYSVDEE--YLTGKAIKLVKGDNQKLIPKTI 1044
QY 570 NNNKYYQ-----SLIGTQWHEBDLVDIIRMEDK-EVLPVTHNLTLRKTVTGLADGRTRDF 623
DB 1045 DHSRLQYKDSSTIYVGDSSMKPEE--NFVSATDKGQDVPE-----KITVSGQVONXXAGV 1097
QY 624 HFEIELKNNKQELLSSQVVKDKTNLEFKDKATINLKHGSLTLQGLPEGSILYVKETDS 683
DB 1098 YPIIYSDEGKEETAYVVKPDQSKLEVKD--TTIYV--GDSWK---PE--DNFVSATDK 1147
QY 684 EGYKV---KVNSEVANATVSKTGITSDETLAFEN 715
DB 1148 TGDVPEPEKIDVQGTVN--VDKIG---DYELVYKN 1177

RESULT 10
ABP43429
ID ABP43429 standard; Protein; 1448 AA.
AC ABP43429;
XX
XX 05-AUG-2002 (first entry)
XX
XX E faecalis EF104 protein.
DE
XX Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.
KW
XX Enterococcus faecalis.
OS
XX
XX US2002045737-A1.
PN
XX
XX 18-APR-2002.
PD
XX
XX 04-MAY-1998; 98US-0071035.
PE
XX
XX 04-MAY-1998; 98US-0071035.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Choi GH, Bailey C, Hromockyj A, Kunsch CA;
PI

```



```

XX  WPI: 2002-425450/45.
DR  N-PSDB; ABN98185.
XX
XX  New genes and polypeptides from Enterococcus faecalis, useful as
XX  vaccines for preventing, treating or attenuating an infection caused by
XX  a member of the Enterococcus genus in an animal, particularly E.
XX  faecalis
XX
XX  Claim 9; Page 182-183; 255pp; English.
XX
XX  The present invention provides the protein and coding sequences of a
XX  number of polypeptides from Enterococcus faecalis. The proteins can be
XX  used as vaccines for preventing or attenuating an infection caused by a
XX  member of the Enterococcus genus in an animal, particularly E. faecalis.
XX  The polynucleotide is also useful for preventing or treating E. faecalis
XX  infection. The present sequence is a protein of the invention.
XX
SQ  Sequence 1448 AA;

Query Match 4.1%; Score 161; DB 23; Length 1448;
Best Local Similarity 21.1%; Pred. No. 0.017;
Matches 159; Conservative 104; Mismatches 326; Indels 166; Gaps 40;

OY 39 ALVTSMGAKTVFGLVES---STPNAINPDSSSEYRWYGYESYVNGHPHYQOFVRAHDLR 95
DB 511 SLSTPVIQPKAIQLVSDQYIEPISVNP-LNAETAMGNVDO-----NCAVSSR 558
OY 96 --VNLEGRSYOVYCFNLKKAFF---LGSDSYKWKYKKHKGISTKFEDYAMSFRITGD 149
DB 559 TTIVSMGSKKEPIQLNLEIKVHPNPLSLRAIKREIFYKK---LGT---DVTVPITSDGS 611
OY 150 ELNOKLRAVYNGHPQANANGIMEGLE-----PLNAIRVTOEAVWYYSNAPISNDESFK 204
DB 612 VIKFTPTTNEIQIPIGFNVVPSLPKDKSIPVDPIPTMSAEGILTPTVTT---NSK 668
OY 205 RESENLTSTQSLMRQALQOLIDPNLATKMPKQVDPDFQSLPESSEDKGPKYKQYON 264
DB 669 RGSERTLOSSKNQPLVNAKRNDSFDSLSVTKIPAGA--DVLEPDIVDSN--DQVDSITPQ 724
OY 265 LLSGLVPTKPPPTGDDPMP--PNOPQTSVLI-----RKYAIGDYSKL---LEGATLQ 313
DB 725 YMDRGQYEDKPMTPSPGYPITFEDENSTYTFDGGKTNKRITII-EYKANGMIDVPTLY 783
OY 314 LTG-----DNVNSFQARVFNSSNDIGERIELSDGTYITELNSPAGISIAEPIITPKVAG 367
DB 784 ITTGAKEPQSNNSGASVYON--EALDILISAT-----QANPTLKNVTKT 828
OY 368 KVTY-IIDGK--QIENPKIEIPEYSAVYNDFEESVLTQNVAKFYAKKNKNGSSQVY 424
DB 829 TVTTKNIDKNTHRKNPTELTPKGTNAQIDLSITY-----KGVPEDA 873
OY 425 YCFNADLKSPDSEDEGKTMPTPTTGE--VKYTHIAGRLFKYTVKPRDTPDTFLKH 481
DB 874 Y-----SLEKTINGAKVIFKDYLTENITIEYTVWSANAGQIYETITSETLNMOSA 926
OY 482 IKKYI-----EKGYREKGAIEYSGLEQLERATQALAIYFPD---SREL 524
DB 927 SKKVVTAAPITLAKSEGDAGIYVLAITFTTHNVEDNQIAKVSFELIDNVTHTATEF 986
OY 525 DKDKLDVHGFGDM-NDSTLAVAKILVEYAODSNPOLT-----DLDFEIP----- 569
DB 987 TTDERGOVSFPAIMTGDTLAVTNVPOEYVDEE--YLTGKAIKLYKDNOLKIPLTITI 1044
OY 570 NNNKTY-----SLIGTOMHPEDLVADIIMEDKK-EVIRPTNHLTKRTVTGLAGDRTKDF 623
DB 1045 DHSRLQVSDSTIYVGDSSWKPEE--NFVSATDKTGQDVPFE-----KITVSGQVNDXKAGV 1097
OY 624 HFEIELKNNKQELLSSQYFKTDKINLEFKDGKATINLKGESLTLQLEGGSIYLYKEDS 683
DB 1098 YPIIYSDEGKEKEIYVYVKPQSSKLEAVD--TTIIV--GDSWK---PE--DNFVSATDK 1147
OY 684 EGYRV---KVNSEVANATVSKTGITSDETLAFEN 715

```

```

DB 1148 TGQDVPEFKIDVQGTIV--VDKIS---DYELIYKN 1177

RESULT 11
ABU13708
ID ABU13708 standard; Protein; 1448 AA.
XX
XX  ABU13708;
XX
XX  26-FEB-2003 (first entry)
XX
XX  Enterococcus faecalis EF040 polypeptide #201.
XX
XX  EF040; immunostimulant; antibacterial; gene mapping.
XX
XX  Enterococcus faecalis.
XX
XX  US6448043-B1.
XX
XX  10-SEP-2002.
XX
XX  04-MAY-1998; 98US-0071035.
XX
XX  06-MAY-1997; 97US-044031P.
XX  16-MAY-1997; 97US-046655P.
XX  14-NOV-1997; 97US-066009P.
XX  14-NOV-1997; 97US-066009P.
XX
XX  (HUMA-) HUMAN GENOME SCI INC.
XX
XX  Chol GH, Bailey C, Hromockyj A, Kunsch CA;
XX
XX  WPI: 2003-089120/08.
XX  N-PSDB; ABX61755.
XX
XX  New EF040 polypeptides and polynucleotides from Enterococcus faecalis,
XX  useful for generating an immune response against E. faecalis and other
XX  Enterococcus species, and as vaccines against other bacterial genera
XX
XX  Example 1; Column 205-206; 146pp; English.
XX
XX  The invention relates to polynucleotide fragments of a gene from
XX  Enterococcus faecalis, EF040, and the polypeptides encoded by them. The
XX  polypeptides are useful in detecting E. faecalis, as epitope tags, as
XX  molecular weight markers on SDS-PAGE gels or for molecular sieve gel
XX  filtration columns, in generating antibodies that specifically bind to
XX  the E. faecalis polypeptides, in generating an immune response against E.
XX  faecalis and other Enterococcus species and as vaccines against other
XX  bacterial genera. The polynucleotides are useful as probes for gene
XX  mapping and for identifying E. faecalis in biological samples. Sequences
XX  CC ABU13508-ABU13755 represent EF040 polypeptides of the invention.
XX  CC Note: The sequence data for this patent can also be obtained from USPTO
XX  at segdata.uspto.gov/sequence.html.
XX
SQ  Sequence 1448 AA;

Query Match 4.1%; Score 161; DB 24; Length 1448;
Best Local Similarity 21.1%; Pred. No. 0.017;
Matches 159; Conservative 104; Mismatches 326; Indels 166; Gaps 40;

OY 39 ALVTSMGAKTVFGLVES---STPNAINPDSSSEYRWYGYESYVNGHPHYQOFVRAHDLR 95
DB 511 SLSTPVIQPKAIQLVSDQYIEPISVNP-LNAETAMGNVDO-----NCAVSSR 558
OY 96 --VNLEGRSYOVYCFNLKKAFF---LGSDSYKWKYKKHKGISTKFEDYAMSFRITGD 149
DB 559 TTIVSMGSKKEPIQLNLEIKVHPNPLSLRAIKREIFYKK---LGT---DVTVPITSDGS 611
OY 150 ELNOKLRAVYNGHPQANANGIMEGLE-----PLNAIRVTOEAVWYYSNAPISNDESFK 204
DB 612 VIKFTPTTNEIQIPIGFNVVPSLPKDKSIPVDPIPTMSAEGILTPTVTT---NSK 668

```


Db 1297 -LTKLDSKSRNLAGEFEELQTKLGSLKDKVYT-EANGQLQIDNLAPGQYLVETKAPT 1354
 QY 551 EYADQSNPQLDLDFEIPNNKKYQ-----SLIGQWHP-----DLV 588
 Db 1355 GVDLADP-----VEFTLEFNOKAPIQVTKTNFMSTGSAVLTKGTGKALLANATFKLV 1409
 QY 589 DIIRMEDKKEVIVPTNHLNLT-----RKVTGLA-GD-----RTK-----DEHFEIELKN 631
 Db 1410 D-----EDNN-----VTENLFTDASGKLEITNLAPGYQLETKAPAGYELDTVPVVKITF 1461
 QY 632 NKQELLS-----QTVKTDKTNLEFKDKGKATINLKHGSLVLQGLPEGYSYLVKETSDEG 685
 Db 1462 DQKETLQVTKTNLKTIVSGKVAEFDVDTKGNV-LAEKEIHT--GI-VGDKYATKAKADIKG 1517
 QY 686 KKV-----KNSQEVANATVSKTGITSDFLAFENKKEPV----- 721
 Db 1518 YKLTQPTNKEGVEFKETDQK-----TFVEKNKAPIVVNDKPVTPVKPTKVPDPACK 1571
 QY 722 -----PTGVDQKINGYALALIVAGISLGIVGHTIRIK 755
 Db 1572 PTVKTSLPSTGDESPYG-----IIFGLFASFWGLFLKRSK 1608
 RESULT 13
 ID ABB47329 standard; Protein; 1530 AA.
 AC ABB47329;
 DT 05-FEB-2002 (first entry)
 DE Listeria monocytogenes protein #33.
 XX Listeria monocytogenes protein #33.
 KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KM vitamin B12; bacterial infection; disease.
 OS Listeria monocytogenes.
 PN WO200177335-A2.
 PD 18-OCT-2001.
 PF 11-APR-2001; 2001WO-FR01118.
 PR 11-APR-2000; 2000FR-0004629.
 XX (INSP) INST PASTEUR.
 PA Buchrieser C, Frangeul L, Couve E, Rusnlok C, Fsihi H, Dehoux P;
 PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Pierrez-Martinez A, Amend A;
 PI Chakraborty T, Dommann E, Hahn T, Berche P, Chablit A, Durrant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Madeno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 DR WPI: 2002-010914/01.
 XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides
 PS Claim 6; SEQ ID No 34; 192pp; French.
 CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and

CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1530 AA;

Query Match 4.0%; Score 156; DB 23; Length 1530;

Best Local Similarity 20.6%; Pred. NO. 0.045; Indels 316; Gaps 38;

Matches 160; Conservative 78; Mismatches 221; Indels 316; Gaps 38;

QY 193 NAPISNPDESFR--ESESNLVSTQSLMRQALQIDPNLTKAPRQVPDPDSIFE 250
 Db 794 NMPISPGEDLKTGYDESN-----LEFYQKFNELNQSIVIK-----YQTAITL 839
 QY 251 SEDK-----GDKYKGYQNLISGLVPTK-----PPPDPPMPNPQPTTSVLIRKAI 300
 Db 840 TSDFTTAQIGNSVFTGDNITKGETEKTKNIEVKITGD--GTGETGKIILNKYDK 896
 QY 301 GDYSKLLGATLQLGDNVNSFOARFSSNDIGELISDGYTLTLELNSPAGYSI---- 356
 Db 897 ADPSIPLGATFDLIA--DDEKVDYDTTKNGVIEFDLDVYGDYTLKEVASAPGYTLPTAS 955
 QY 357 AEPITKVE-----AGKVTYIID--GKQLEN 380
 Db 956 TENIQVKLEQDEKVVQVWNEKPIKETGEVHLVTKDNTGATLAAEESIVDKSAGELON 1015
 QY 381 -----PKKEIV-----EPYSVAIYDFE--EFSVLTQNTAKFYAKN--- 417
 Db 1016 GLTIDENGLTIHNLDSYLLKTKRABEGYKLSKETEFSVSGVDALIEQANERDL 1075
 QY 418 -----NGSSQV-----YCF----- 427
 Db 1076 GEAVLTQKVDSETNAKLSGAKFNLNDSGEVIOTNLVSDENGIRVQNLPEGDYAFQETEA 1135
 QY 428 --NADLKS-----PPDSBDG-----GKT----- 443
 Db 1136 PTNYDLATNTWFTLIVAGOTSATWTAENNKTKGPDVDTGEVILVKKDSAGETLEGAVF 1195
 QY 444 --MTPD-----FTT-----GEVYTHIA-----GROLFKYTVAPROTD 474
 Db 1196 DLMTADGAIVASNLTTDANGELITVNLAPGKYSFKETRAPEGYELATVWETIAPNQPE 1255
 QY 475 -----PDT--FLKHKKVIRKGYREKQALEYSGLTQLRATQALAIYFT 519
 Db 1256 KITTAENMTKLAPIPDASVKTIKQDSENGVRLAG--AEFSLIAENGELQTNLK-----T 1309
 QY 520 DSAELDKDLKLDYHGFQMDNSTLAVAKILVEYADQSNPQITDLDFIPINNKKYQSLIG 579
 Db 1310 DEA-----GELEVNNLAPGNVRI--QETKAP-----DGYQ--LES 1340
 QY 580 TQWHPEDLVDIRMEDKKEVIVPTNHLNLTQKTY--TGLADRTKPFHEIELKNNKQEL 637
 Db 1341 TPWQFE-----IVANDTSQVTVIAENALPEPVAETGAVRLIKTQTSSETGRLSGAVSLL 1395
 QY 638 SOTVTKDTNLEFKGKATINLKHGSLVLQGLPEGYSYLVKETS--GGYK----- 688
 Db 1396 DESGVIANLTTDE-----NKE--IFIDSLTIG--NYSLEKTAAPGCGYELAEQPMWQ 1445
 QY 689 -----KNSQEVANATVSKTGITSD-----ETLAFENKKEP 719
 Db 1446 IVKQGVDAVVIKAENSPITANGAISFEQEDTKPESIEIPVKTTTLATVTKLP 1500
 RESULT 14
 ID ABB47336 standard; Protein; 832 AA.

XX AC ABB47336;
 XX DE 05-FEB-2002 (first entry)
 XX DE Listeria monocytogenes protein #40.
 XX KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 XX KM vitamin B12; bacterial infection; disease.
 XX OS Listeria monocytogenes.
 XX PN W0200177335-A2.
 XX PD 18-OCT-2001.
 XX PF 11-APR-2001; 2001WO-FR011118.
 XX PR 11-APR-2000; 2000FR-0004629.
 XX PA (INSP) INST PASTEUR.
 XX PI Buchrieser C, Frangoul L, Couve E, Rusniok C, Eshti H, Dehoux P,
 PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P,
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA,
 PI Dominguez-Bernal G, Garrido-Garcia P, Tlerriz-Martinez A, Amend A,
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L,
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
 PI Menduño E, De Pablos B, Wehlend J, Kaerst U, Entian K, Hauf J,
 PI Rose M, Voss H.
 XX DR WPI; 2002-010914/01.
 XX PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides
 XX PS Claim 6; SEQ ID No 41; 192pp; French.
 XX CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes BGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by L.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 832 AA;
 Query Match 3.9%; Score 155; DB 23; Length 832;
 Best Local Similarity 19.9%; Pred. No. 0.02;
 Matches 141; Conservative 93; Mismatches 232; Indels 244; Gaps 33;
 QY 155 LRAVMNGHPQNNAN-GIMEGLEPLNARVQEAWVWYSDNAPISNPDESKRESESLV- 212
 DB 176 LKSIANNKTTGNFSLVKTLPPLHTLV-----LGNATTELDIENQPLVLT 222
 QY 213 -STSOISIMROALKOLIDPN-----LATKMPKOVDPDDQLSIFESE 252
 DB 223 LSNDELLEKTLTKLNLSQNLGLRIASSISIDMGDLESYVLMNLPETISYDISGNVYDSD 282
 QY 253 DKGDYKNGYONLISGLVFTKPPTPGDDPPMPNPQOTSVLIRKKAIG--DYSKLL-- 307

DB 283 DIHLENPAAVKNNDISSNELTRLPKINDPFL-----LTTINRSKNIDRLSSKILVDVP 336
 QY 308 -----EGATLQLT-----GDNVNSFOA-----RVFSNDI-----GER 335
 DB 337 KLATLNADKQAVTLSTKITAAGNFTIPNNVNLNLAGQWVTPKLIISNGYSQSIAMASGEL 396
 QY 336 IELSDGYTLTE-LNSPA-----GYSIAEPTFEKVEAGKYVYTIIDKQLE----- 379
 DB 397 SGLSKVSYTFDEYVINSALIGKXTGTYNQPIEYKA-----VPYVAKSVSYAPVANKDEA 452
 QY 380 -----NPKELVEPYSVBAYNDEEFESVLTQNAKFEYVANKNGSSOVVYCFNA 429
 DB 453 TFLQDIRASASENAQITSDYS-----EVYDAFPGDYVTVLHAKNE-----F 494
 QY 430 DLKSPDSEDCGKTMPDFTTGEVKYTHIAGRLDFKYVYPRRTDDPTFLKHIKYYENG 489
 DB 495 DLKA-----DPYTVVYHINDIQKPO 514
 QY 490 YREKGAIEY---SGLTETQLRA-----ATQL---AIYPTDSAEIDKDKLDYHG--- 534
 DB 515 VAVNSNDISFEVGETELTSEVLAKSGAVYTDLDEAIKMEVDLSBVDSSKLTGYEATIA 574
 QY 535 ---FGDMNSTLAVAKILVEYADSNPP--QLTDLDFIIPNNKKYQ-----SLIGQ 561
 DB 575 KSKSGASSDPDKLSVKIV-----DTEKPIQIINNPELIEKSELTGQIIDQGITATD 629
 QY 582 WHREDI---VDIIMEDKKFVIVTNHLTKRTVYGLAGRTDPHFETELKNNKELL 638
 DB 630 NYDODLNIHMDLSKVDKSR---PGSYEVNT-----YTEDSSGKSEYVY 670
 QY 639 QTVKTDKTNLEFKDKRATI-----NLKHGESLTLQG-LPEGSYLVKRETSRGYKVVN 691
 DB 671 ITVAKVEPARI---GRTTIYMOSENNELAESNTITGEVETETLAKEL--EGYTLKEN 724
 QY 692 SQ-----EVANATVS---KTGITSDETLAEFNKKEPVYPTGVQDKING 731
 DB 725 PANSGVFEETROTQIYIVKDIINPEFVYSNNVTPELPSNNNSVNG 774
 RESULT 15
 AAM17900
 ID AAM17900 standard; Protein: 1849 AA.
 XX AAM17900;
 XX 29-JAN-1998 (first entry)
 XX DE Photorhabdus luminescens insect toxin TcdA11.
 XX KW Insecticide; insect; toxin; pest control; biological control;
 KW Photorhabdus luminescens; TcdA; Southern corn rootworm;
 KW Colorado potato beetle; Western corn rootworm; meal worm;
 KW boll weevil; turf grub; Coleoptera; beet armyworm; black cutworm;
 KW cabbage looper; codling moth; corn earworm; European corn borer;
 KW tobacco hornworm; tobacco budworm; Lepidoptera; Hymenoptera;
 KW Diptera; Dictyoptera; Acarina; Homoptera.
 XX OS Photorhabdus luminescens strain W-14 (ATCC 55397).
 XX FH Key Location/Qualifiers
 FT Protein 1..1849
 FT /label= TcdA11
 FT Peptide 1..12
 FT /note= "S2 N-terminus (Claim 30)"
 FT Peptide 196..211
 FT /note= "tryptic peptide (Claim 30)"
 FT Peptide 466..475
 FT /note= "tryptic peptide (Claim 30)"
 FT Peptide 993..1004
 FT /note= "isolated N-terminal peptide (Claim 30)"
 FT Peptide 1297..1312
 FT /note= "tryptic peptide (Claim 30)"

